

Laplacian Pyramid Reconstruction and Refinement for Semantic Segmentation

Golnaz Ghiasi and Charless C. Fowlkes
{gghiasi,fowlkes}@ics.uci.edu

Dept. of Computer Science,
University of California, Irvine

ECCV 2016

目的:

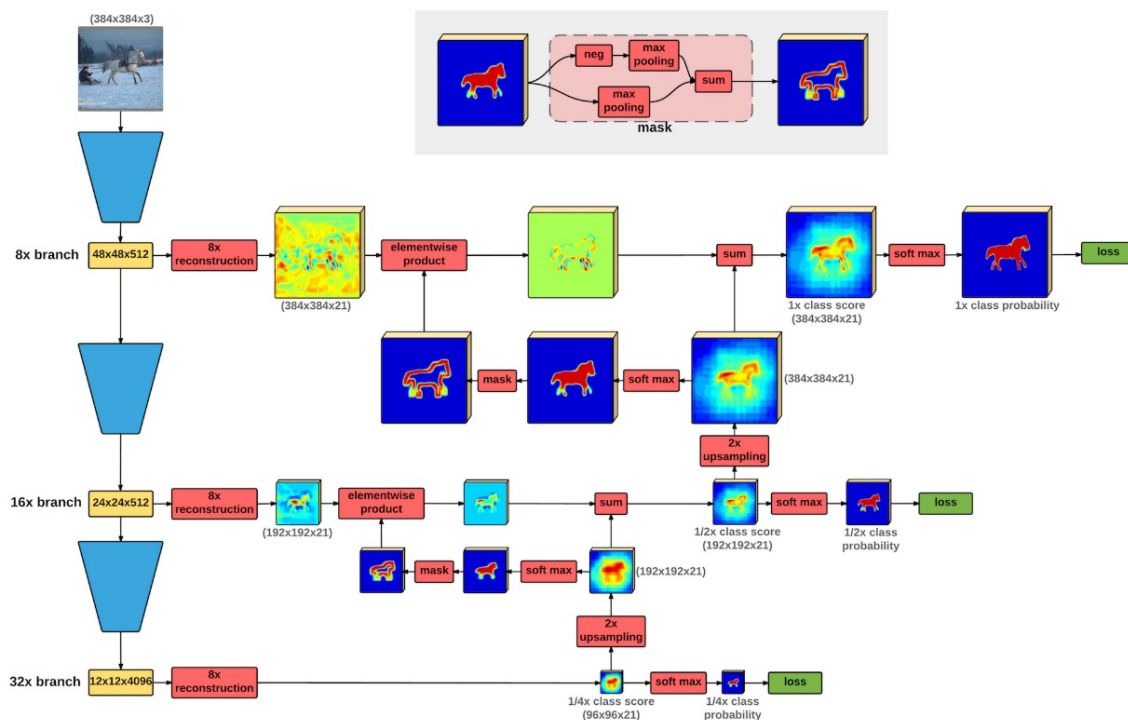
使用 multiplicative gating 来更有效地整合底层位置信息和高层语义信息。

其思想来源于拉普拉斯金字塔。

方法:

随着网络层数加深，深层特征图的语义信息丰富，但是分辨率太小，缺少高频的边界信息，高频成分越来越少。如果直接相加，浅层特征图的噪声会对原本深层特征图中置信度较高的点产生干扰，因此网络会减少浅层特征图的权重，对预测的贡献很小。

所以本文提出的方法，去掉了高频特征图中的低频部分，只留下边界信息，然后去 refine 低分辨率的特征图。在低分辨率预测对语义内容有信心的区域，使用了乘法门控，避免了集成高分辨率输出后出现的噪声。



总结：

使用池化而不是卷积来生成边界的 mask，没有加入额外的参数量，使用 mask 也类似于加入了注意力机制，告诉网络每一层的预测结果哪些重要哪些不重要，再利用相加把这层的影响叠加到上一层的分割上，使深层的分割更多得决定分类，浅层的分割更多得决定边缘，得到更精确的分割结果。不过也可以直接增加一个卷积层，让网络自己学习如何去除噪声或实现其他变换，类似 refinenet 这种。